Finite Dimesional Spaces Vol.II
by Walter Noll (∼1990)

2. Differential Calculus II

by Walter Noll, 2007

In this chapter, the phrase "let... be a linear space" will be used as a shorthand for "let... be a finite-dimensional linear space over $\mathbb{R}$". However, many definitions remain meaningful and many results remain valid when some or all of the given spaces fail to be finite-dimensional.

21.Exterior Differentials.

In this section, we assume that a flat space $\mathcal{E}$ with translation space $\mathcal{V}$ and an open subset $\mathcal{D}$ of $\mathcal{E}$ are given. All the results presented here depend on the Curl-Gadient Theorem of Sect.611 in Vol.I., but they involve some complex book-keeping.

First, we assume that a linear space $\mathcal{V}'$ and a mapping

$$H : \mathcal{D} \longrightarrow \operatorname{Lin}(\mathcal{V}, \mathcal{V}') \tag{21.1}$$

are given. If H is differentiable, we can define

$$\operatorname{Curl} H : \mathcal{D} \longrightarrow \operatorname{Skew}_2(\mathcal{V}^2, \mathcal{V}') \tag{21.2}$$

according to (611) in Vol.I, so that

$$\operatorname{Curl} H(\mathbf{u}, \mathbf{v}) = (\nabla H\mathbf{u})\mathbf{v} - (\nabla H\mathbf{v})\mathbf{u} \qquad \text{for all} \quad \mathbf{u}, \mathbf{v} \in \mathcal{V}\ . \tag{21.3}$$

We understand this and the following equations be taken value-wise, so that, for example $\operatorname{Curl} H(\mathbf{u}, \mathbf{v})(x) = \operatorname{Curl} H(x)(\mathbf{u}, \mathbf{v})$ for all $x \in \mathcal{D}$.

We will use the following notations:

For every $\mathbf{v} \in \mathcal{V}$ and every list $\mathbf{f} \in \mathcal{V}^k$ we define $\mathbf{f}.\mathbf{v} \in \mathcal{V}^{(k+1)}$ by

$$(\mathbf{f}.\mathbf{v})_i := \begin{cases} \mathbf{f}_i & \text{if } i \in k^] \\ \mathbf{v} & \text{if } i = k+1 \end{cases} \ . \tag{21.4}$$

In words, the list $\mathbf{f}.\mathbf{v}$ is obtained from the list $\mathbf{f}$ by attaching $\mathbf{v}$ at the end. (This is a special case of concatenation as defined by (02.13).)

Given $k \in \mathbb{N}^\times$ and a mapping

$$W : \mathcal{D} \longrightarrow \operatorname{Skew}_{k+1}(\mathcal{V}^{k+1}, \mathcal{V}')\ , \tag{21.5}$$

we define

$$W_* : \mathcal{D} \longrightarrow \operatorname{Lin}(\mathcal{V}, \operatorname{Skew}_k(\mathcal{V}^k, \mathcal{V}')) \tag{21.6}$$

by

$$W_*(\mathbf{v})(\mathbf{f}) := W(\mathbf{f}.\mathbf{v}) \qquad \text{for all} \quad \mathbf{v} \in \mathcal{V},\ \mathbf{f} \in \mathcal{V}^k \ , \tag{21.7}$$

which, again, is to be understood value-wise.

Proposition 1. *Assume that the mapping* H *as described by (21.1) is twice differentiable. Then* $\nabla(\mathrm{Curl}\, H)_* \in \mathrm{Lin}\,(\mathcal{V}, \mathrm{Lin}\,(\mathcal{V}, \mathrm{Lin}\,(\mathcal{V}, \mathcal{V}')))$ *satisfies*

$$\begin{aligned}&(((\nabla(\mathrm{Curl}\, H)_*\mathbf{v}_1)\mathbf{v}_2)\mathbf{v}_3) + (((\nabla(\mathrm{Curl}\, H)_*\mathbf{v}_2)\mathbf{v}_3)\mathbf{v}_1)\\ &+(((\nabla(\mathrm{Curl}\, H)_*\mathbf{v}_3)\mathbf{v}_1)\mathbf{v}_2) = 0 \qquad \text{for all} \quad \mathbf{v} \in \mathcal{V}^3 \ .\end{aligned} \tag{21.8}$$

Note that the last two terms of the sum in (21.8) are obtained from the first by a cyclic permutation of (1,2,3).

Proof: Let a triple $\mathbf{v} \in \mathcal{V}^3$ be given. Differentiation of (21.3) gives

$$(((\nabla(\mathrm{Curl}\, H)_*\mathbf{v}_1)\mathbf{v}_2)\mathbf{v}_3) = (((\nabla^{(2)}H\mathbf{v}_1)\mathbf{v}_2)\mathbf{v}_3) - (((\nabla^{(2)}H\mathbf{v}_1)\mathbf{v}_3)\mathbf{v}_2) \ . \tag{21.9}$$

We rewrite this equation with (1,2,3) replaced by (2,3,1) and then by (3,1,2) and add the three equations thus obtained. We find that the result is a sum of 6 terms. Combining the first with the forth term of this sum, we get $(((\nabla^{(2)}H\mathbf{v}_1)\mathbf{v}_2)\mathbf{v}_3) - (((\nabla^{(2)}H\mathbf{v}_2)\mathbf{v}_1)\mathbf{v}_3)$, which is zero by the Theorem on Symmetry of Second Gradients in Sect.611 of Vol.I (a restatement of the first part of the Curl-Gradient Theorem). Similarly, combining the second with the fifth term and combining the third with the sixth give zero also, which proves that the left side of (21.8) is zero. ∎

Proposition 2. *Assume that the mapping* H *as described by (21.1) is twice differentiable. Then* $\mathrm{Curl}\,(\mathrm{Curl}\, H)_* \in \mathrm{Skew}_{\,2}(\mathcal{V}^2, \mathrm{Lin}\,(\mathcal{V}, \mathcal{V}')))$ *satisfies*

$$\begin{aligned}&((\mathrm{Curl}\,(\mathrm{Curl}\, H)_*(\mathbf{v}_1, \mathbf{v}_2)\mathbf{v}_3) + ((\mathrm{Curl}\,(\mathrm{Curl}\, H)_*(\mathbf{v}_2, \mathbf{v}_3)\mathbf{v}_1)\\ &+((\mathrm{Curl}\,(\mathrm{Curl}\, H)_*(\mathbf{v}_3, \mathbf{v}_1)\mathbf{v}_2) = 0 \qquad \text{for all} \quad \mathbf{v} \in \mathcal{V}^3 \ .\end{aligned} \tag{21.10}$$

Again, the last two terms of the sum in (21.10) are obtained from the first by a cyclic permutation of (1,2,3)

Proof: Let a triple $\mathbf{v} \in \mathcal{V}^3$ be given. We write down the equation (21.8) and subtract from it the equation obtained from (21.8) by replacing (1,2,3) by (2,1,3). Using (21.3), it is easily seen that (21.10) is valid. ∎

Definition 1: *Let* $k \in \mathbb{N}^\times$ *be given. A mapping* $\omega : \mathcal{D} \to \mathrm{Skew}_{\,k}(\mathcal{V}^k, \mathbb{R})$ *is called a* **skew** k**-form**.

We define $\omega_* : \mathcal{D} \longrightarrow \mathrm{Lin}\,(\mathcal{V}, \mathrm{Skew}_{\,k-1}(\mathcal{V}^{k-1}, \mathbb{R}))$ according to (21.6), so that, observing (02.1), wed have

$$\omega_*(\mathbf{f}_k)(\mathbf{f}_{\underline{)k(}}) = \omega(\mathbf{f}) \qquad \text{for all} \quad \mathbf{f} \in \mathcal{V}^k \ . \tag{21.11}$$

We have

$$\omega_*(\mathbf{f}_j)(\mathbf{f}_{\underline{)j(}}) = (-1)^{(k-j)}\omega(\mathbf{f}) \qquad \text{for all} \quad j \in k^] \ \text{ and } \ \mathbf{f} \in \mathcal{V}^k \tag{21.12}$$

because the list $\mathbf{f}_{\underline{)j(}}.\mathbf{f}_j$ differs from the list $\mathbf{f}$ by $k-j$ switches.

If ω is a differentiable skew k-form and if ω_* is defined according to (21.11), then Curl ω_*, as defined by (21.3), has values in Skew $_2(\mathcal{V}^2, \text{Skew}_{(k-1)}(\mathcal{V}^{(k-1)}, \mathbb{R}))$ and is given by

$$(\text{Curl } \omega_*(\mathbf{u},\mathbf{v}))(\mathbf{h}) = (((\nabla\omega_*)\mathbf{u})\mathbf{v})(\mathbf{h}) - (((\nabla\omega_*)\mathbf{v})\mathbf{u})(\mathbf{h})$$

$$\text{for all} \quad \mathbf{u},\mathbf{v} \in \mathcal{V}, \ \ \mathbf{h} \in \mathcal{V}^{k-1} \ . \tag{21,13}$$

Proposition 3: *With every differentiable skew $k-$form ω we can then associate skew $(k+1)$-form, denoted* Curl $_*\omega$, *which satisfies (see(02.2))*

$$\text{Curl}_*\omega(\mathbf{g}) := \text{Curl } \omega_*(\mathbf{g}_p, \mathbf{g}_{p+1})(\mathbf{g}_{\underline{)p,p+1(}}) \quad \text{for all} \ \ p \in (k-1)^] \ \text{ and } \mathbf{g} \in \mathcal{V}^{k+1} \ . \tag{21.14}$$

Proof: Let a differentiable skew $k-$form ω and $\mathbf{g} \in \mathcal{V}^{k+1}$ be given. Using (21.13) with the choice $\mathbf{f} := \mathbf{g}_{\underline{)k+1(}}$ and hence $\mathbf{f}_{\underline{)k(}} := \mathbf{g}_{\underline{)k,k+1(}}$ gives $\omega_*(\mathbf{g}_k)(\mathbf{g}_{\underline{)k,k+1(}}) = \omega(\mathbf{g}_{\underline{)k+1(}})$.Taking the gradient in the direction of $\mathbf{g}_{k+1}$ and using the fact that $\mathbf{g}_{\underline{)k,k+1(}}.\mathbf{g}_k = \mathbf{g}_{\underline{)k+1(}}$, we find that

$$(((\nabla\omega_*)\mathbf{g}_{k+1})\mathbf{g}_k)(\mathbf{g}_{\underline{)k,k+1(}}) = ((\nabla\omega)\mathbf{g}_{(k+1)})(\mathbf{g}_{\underline{)k+1(}}) \ . \tag{21.15}$$

Now, it is clear that the left side of (21.15) switches sign if any two of the terms of $\mathbf{g}_{\underline{)k,k+1(}}$, i.e., any two of the first $k-1$ terms of $\mathbf{g}$, are switched. The right side of (21.15) switches sign if the last two terms of $\mathbf{g}_{\underline{)k+1(}}$, which are the k^{th} and the $(k-1)^{th}$ term of $\mathbf{g}$ itself, are switched. Now put

$$\text{Curl}_*\omega(\mathbf{g}) := \text{Curl } \omega_*((\mathbf{g}_k, \mathbf{g}_{k+1})(\mathbf{g}_{\underline{)k,k+1(}}) \ . \tag{21.16}$$

It follows also that this value switches sign if any two of the first $k-1$ terms and the k^{th} and the $(k+1)^{th}$ term of $\mathbf{g}$ are switched. It is clear from the fact that the values of a Curl are skew that (21.15) also switches sign if the last two terms of $\mathbf{g}$ are switched. By the characterization of skew mappings given by Prop.4 of Sect.01 it follows that Curl $_*\omega$ has skew values. Of course, its values are also multilinear and hence it is a skew (k+1)-form.

Now, given $p \in k^]$, it is easily seen that the list $\mathbf{g}_{\underline{)p,p+1(}} .\mathbf{g}_p.\mathbf{g}_{p+1}$ differs from the list $\mathbf{g}$ by $2(k-p)$ switches and hence an even permutation. Therefore (21.14) follows from (21.15) and the fact that Curl $_*\omega$ has skew values. ∎

Definition 2: *We denote by SKC_k^m the set of all skew k-forms that are of class C^m.*

Let $\omega \in SKC_k^m$ *and* $\mathbf{g} \in \mathcal{V}^{(k+1)}$ *be given. We define* $\mathbf{d}\omega(\mathbf{g}) : \mathcal{D} \to I\!R$ *by*

$$\mathbf{d}\omega(\mathbf{g}) := \Lambda(\nabla\omega)(\mathbf{g}) = \sum_{j \in (k+1)]} (-1)^{k+1-j}((\nabla\omega)\mathbf{g}_j)(\mathbf{g}_{\underline{)j(}}) \ , \tag{21.17}$$

where the definition (11.3) of Λ *in Sect.11 is used. By the Lemma in Sect.11, the values of* $\Lambda(\nabla\omega)$ *belong to* $\text{Skew}_{(k+1)}(\mathcal{V}^{(k+1)}, I\!R)$. *Hence* $\mathbf{d}\omega : \mathcal{D} \to \text{Skew}_{(k+1)}(\mathcal{V}^{(k+1)}, I\!R)$ *may and will be considered as a member of* $SKC_{(k+1)}^{(m-1)}$. *The skew* $k+1$ *form* $\mathbf{d}\omega$ *thus obtained is called the* **exterior differential** *of the skew* k *form* ω.

Proposition 4 : *Given* $\omega \in SKC_k^m$ *, we have*

$$2\mathbf{d}\omega = (k+1)\text{Curl}_*\omega \ , \tag{21.18}$$

where $\text{Curl}_*\omega$ *is defined according Prop.3.*

Proof: Let $\mathbf{g} \in \mathcal{V}^{k+1}$ and $j \in k]$ be given. Using (21.14) with the choice $\mathbf{f} := \mathbf{g}_{\underline{)j(}}$, we find

$$\omega(\mathbf{g}_{\underline{)j(}}) = (-1)^{k-j}\omega_*(\mathbf{g}_{j+1})(\mathbf{g}_{\underline{)j,j+1(}}) \ .$$

Taking the gradient gives

$$((\nabla\omega)\mathbf{u})(\mathbf{g}_{\underline{)j(}}) = (-1)^{k-j}(((\nabla\omega_*)\mathbf{u})\mathbf{g}_{j+1})(\mathbf{g}_{\underline{)j,j+1(}}) \qquad \text{for all} \quad \mathbf{u} \in \mathcal{V} \ .$$

Using this formula with $\mathbf{u} := \mathbf{g}_j$ we obtain

$$((\nabla\omega)\mathbf{g}_j)(\mathbf{g}_{\underline{)j(}}) = (-1)^{k-j}(((\nabla\omega_*)\mathbf{g}_j)\mathbf{g}_{j+1})(\mathbf{g}_{\underline{)j,j+1(}}) \ . \tag{21.19}$$

Similarly, using (21.12) with the choice $\mathbf{f} := \mathbf{g}_{\underline{)j+1(}}$, we obtain

$$((\nabla\omega)\mathbf{g}_{j+1})(\mathbf{g}_{\underline{)j+1(}}) = (-1)^{k-j}(((\nabla\omega_*)\mathbf{g}_{j+1})\mathbf{g}_j)(\mathbf{g}_{\underline{)j,j+1(}}) \ . \tag{21.20}$$

Now, rearranging the summation in (21.17) by putting the first term last, changing the summation index from j to $j+1$ and putting $\mathbf{g}_{k+2} := \mathbf{g}_1$, we obtain

$$\mathbf{d}\omega(\mathbf{g}) = \sum_{j \in (k+1)]} (-1)^{k-j}((\nabla\omega)\mathbf{g}_{j+1})(\mathbf{g}_{\underline{)j+1(}}) \ . \tag{21.21}$$

In view of (21.19), (21,20), and (21.3), we have

$$(-1)^{k-j}\nabla(\omega(\mathbf{g}_{\underline{)j+1(}}))(\mathbf{g}_{j+1}) + (-1)^{k+1-j}\nabla(\omega(\mathbf{g}_{\underline{)j(}}))(\mathbf{g}_j) =$$

$$= \text{Curl}\omega_*(\mathbf{g}_{j+1}, \mathbf{g}_j)(\mathbf{g}_{\underline{)j,j+1(}}) \ .$$

Hence, adding the equations (21.17) and (21.21) and combining the terms corresponding to each $j \in k^{]}$, we find

$$2\mathbf{d}\omega(\mathbf{g}) := \sum_{j \in (k+1)^{]}} \mathrm{Curl}\omega_*(\mathbf{g}_{j+1}, \mathbf{g}_j)(\mathbf{g}_{\underline{)j,j+1(}}) \ . \tag{21.22}$$

The desired result (21.18) follows from (21.14) and (21.22). ∎

Theorem: *Let* $\omega \in \mathrm{SKC}_k^m$ *, with* $k, m \in \mathbb{N}^\times$*, be given. If* $\omega = \mathbf{d}\mu$ *for some* $\mu \in \mathrm{SKC}_{(k-1)}^{(m+1)}$*, then* $\mathbf{d}\omega = 0$*. Conversely, if* $\mathcal{D}$ *is convex and if* $\mathbf{d}\omega = 0$*, then* $\omega = \mathbf{d}\mu$ *for some* $\mu \in \mathrm{SKC}_{(k-1)}^{(m+1)}$*.*

Proof : Assume that $\omega = \mathbf{d}\mu$ for some $\mu \in \mathrm{SKC}_{(k-1)}^{(m+1)}$. Applying Prop.4 to both ω and μ, we find that

$$\mathbf{d}\omega = \frac{k+1}{2}\mathrm{Curl}_*\omega = \frac{(k+1)k}{4}\mathrm{Curl}_*(\mathrm{Curl}_*\mu) \ . \tag{21.23}$$

Now let $\mathbf{h} \in \mathcal{V}^{(k-2)}$, $\mathbf{v} \in \mathcal{V}^3$ and $\mathbf{u} \in \mathcal{V}$ be given. Then $\mathbf{h}.\mathbf{v}_1.\mathbf{u} \in \mathcal{V}^k$. Using Prop.3 with ω replaced by μ, k by $k-1$, $\mathbf{g}$ by $\mathbf{h}.\mathbf{v}_1.\mathbf{v}_2$, and p by $k-2$, we find

$$\mathrm{Curl}_*\mu(\mathbf{h}.\mathbf{v}_1.\mathbf{u}) = \mathrm{Curl}\,\mu_*(\mathbf{v}_1.\mathbf{u})(\mathbf{h}) \ ,$$

and hence, using (21.7) on both sides of this equation, ,

$$((\mathrm{Curl}_*\mu)_*\mathbf{u})(\mathbf{h}.\mathbf{v}_1) = (\mathrm{Curl}\,\mu_*)_*\mathbf{u})\mathbf{v}_1)(\mathbf{h}) \ . \tag{21.24}$$

We note that $(\mathrm{Curl}_*\mu)_* : \mathcal{D} \to \mathrm{Lin}(\mathcal{V}, \mathrm{Skew}_{k-1}(\mathcal{V}^{(k-1)}, \mathbb{R}))$ and hence that $\mathrm{Curl}(\mathrm{Curl}_*\mu) : \mathcal{D} \to \mathrm{Skew}_2(\mathcal{V}^2, \mathrm{Skew}_{k-1}(\mathcal{V}^{(k-1)}, \mathbb{R}))$. . We also note that that $(\mathrm{Curl}\,\mu_*)_* : \mathcal{D} \to \mathrm{Lin}(\mathcal{V}, \mathrm{Lin}(\mathcal{V}, \mathrm{Skew}_{k-2}(\mathcal{V}^{(k-2)}, \mathbb{R})))$ and hence that $\mathrm{Curl}(\mathrm{Curl}\,\mu_*)_* : \mathcal{D} \to \mathrm{Skew}_2(\mathcal{V}^2, \mathrm{Lin}(\mathcal{V}, \mathrm{Skew}_{k-2}(\mathcal{V}^{(k-2)}, \mathbb{R})))$. Since $\mathbf{u} \in \mathcal{V}$ was arbitrary, it follows from (21.24) that

$$\mathrm{Curl}(\mathrm{Curl}_*\mu)_*(\mathbf{v}_2, \mathbf{v}_3)(\mathbf{h}.\mathbf{v}_1) = \mathrm{Curl}(\mathrm{Curl}\,\mu_*)_*(\mathbf{v}_2, \mathbf{v}_3)\mathbf{v}_1)(\mathbf{h}) \ . \tag{21.25}$$

We now use Prop.3 with ω replaced by $\mathrm{Curl}_*\mu$, $\mathbf{g}$ by $\mathbf{h}.\mathbf{v}_1.\mathbf{v}_2.\mathbf{v}_3$, and p by $k-1$, showing that

$$\mathrm{Curl}_*(\mathrm{Curl}_*\mu)(\mathbf{h}.\mathbf{v}_1.\mathbf{v}_2.\mathbf{v}_3) = \mathrm{Curl}(\mathrm{Curl}_*\mu)_*(\mathbf{v}_2, \mathbf{v}_3)(\mathbf{h}.\mathbf{v}_1) \ ,$$

and hence, by (21,27), that

$$\mathrm{Curl}_*(\mathrm{Curl}_*\mu)(\mathbf{h}.\mathbf{v}_1.\mathbf{v}_2.\mathbf{v}_3) = \mathrm{Curl}(\mathrm{Curl}\,\mu_*)_*(\mathbf{v}_2, \mathbf{v}_3)\mathbf{v}_1)(\mathbf{h}) \ . \tag{21.26}$$

Since $\mathrm{Curl}_*(\mathrm{Curl}_*\mu)$ has skew values and since the cyclic permutations of $(1, 2, 3)$ are even, it follows that the right side of (21.26) remains unchanged if $(1, 2, 3)$ is replaced by $(2, 3, 1)$ or $(3, 1, 2)$. We now use Prop.3 with $\mathcal{V}'$ replaced by $\text{Skew}_{(k-2)}(\mathcal{V}^{(k-2)}, \mathbb{R})$ and H by $\mu_* : \mathcal{D} \to \text{Lin}(\mathcal{V}, \text{Skew}_{(k-2)}(\mathcal{V}^{(k-2)}, \mathbb{R})$. It shows that if we replace $(1,2,3)$ by $(2,3,1)$, and $(3,1,2)$ on the right side of (21.26) and add, we obtain zero. Therefore, three times the left hand side of (21.26)is zero, and hence this right hand side itself is zero. Since $\mathbf{h} \in \mathcal{V}^{(k-2)}$ and $\mathbf{v} \in \mathcal{V}^3$ were arbitrary, it follows that $\text{Curl}_*(\text{Curl}_*\mu) = 0$ and hence, by (21.23), that $\mathbf{d}\omega = 0$.

Assume now that $\mathcal{D}$ is convex and that $\mathbf{d}\omega = 0$. By Prop.4, it follows that $\text{Curl}_*\omega = 0$, and hence, by Prop.3, that $\text{Curl}\omega_* = 0$. By the second part of the Curl-Gadient Theorem of Sect.611 in Vol.I, it follows that we can choose a mapping $\nu : \mathcal{D} \to \text{Lin}(\mathcal{V}, \text{Skew}_{k-1}(\mathcal{V}^{k-1}, \mathbb{R}))$ of class C^{m+1} such that

$$\omega_* = \nabla\nu \ . \tag{21.27}$$

Now let $\mathbf{f} \in \mathcal{V}^k$ and $j \in k^]$ be given. It is easily seen that the list $\mathbf{f}_{\underline{)j(}}.\mathbf{f}_j$ differs from $\mathbf{f}$ by $k-j$ switches and hence, using (21.7) and (21.27), that

$$\omega(\mathbf{f}) = (-1)^{k-j}\omega(\mathbf{f}_{\underline{)j(}}.\mathbf{f}_j) = (-1)^{k-j}\omega_*(\mathbf{f}_j)(\mathbf{f}_{\underline{)j(}}) = (-1)^{k-j}((\nabla\nu)\mathbf{f}_j)(\mathbf{f}_{\underline{)j(}}) \ . \tag{21.28}$$

Taking the sum of over all $j \in k^]$ and using Def,2, applied to ν instead of ω, we obtain

$$k\,\omega(\mathbf{f}) = \sum_{j\in k^]}(-1)^{k-j}((\nabla\nu)\mathbf{f}_j)(\mathbf{f}_{\underline{)j(}}) = \mathbf{d}\nu(\mathbf{f}) \ . \tag{21.29}$$

Therefore, since $\mathbf{f} \in \mathcal{V}^k$ was arbitrary, we have $\omega = \mathbf{d}\mu$ with $\mu := \frac{1}{k}\nu$. ▮

22. Higher Gradients

In this section, we assume that open subsets $\mathcal{D}$ and $\mathcal{D}'$ b of flat spaces $\mathcal{E}$ and $\mathcal{E}'$ with tranlation spaces $\mathcal{V}$ and $\mathcal{V}'$ are given.

We extend the Def.1 of Sect 63 of Vol.I as follows

Definition 1. *For $m \in \mathbb{N}^\times$, we say that a given mapping*

$$\phi : \mathcal{D} \longrightarrow \mathcal{D}' \tag{22.1}$$

is m times differentiable and we define its m^{th} **gradient** *$\nabla^{(m)}\phi$ recursively as follows: If ϕ is differentiable, we put*

$$\nabla^{(1)}\phi := \nabla\phi : \mathcal{D} \longrightarrow \text{Lin}(\mathcal{V}, \mathcal{V}') \ . \tag{22.2}$$

If, for a given $m \in \mathbb{N}^\times$, ϕ is m times differentiable and

$$\nabla^{(m)}\phi : \mathcal{D} \longrightarrow \text{Lin}_m(\mathcal{V}^m, \mathcal{V}') \tag{22.3}$$

has been defined and is differentiable, we form its gradient

$$\nabla(\nabla^{(m)}\phi) : \mathcal{D} \longrightarrow \text{Lin}(\mathcal{V}, \text{Lin}_m(\mathcal{V}^m, \mathcal{V}')). \tag{22.4}$$

Using the natural isomorphism (03.9) mediated by (03.10) and (03.11,)we define

$$\nabla^{(m+1)}\phi : \mathcal{D} \longrightarrow \text{Lin}_{\,m+1}(\mathcal{V}^{m+1}, \mathcal{V}')$$

by

$$(\nabla^{(m+1)}\phi)(x) := (\nabla(\nabla^{(m)}\phi)(x))_{<\rightarrow>} \;\in\; \text{Lin}_{\,m+1}(\mathcal{V}^{m+1}, \mathcal{V}'))$$

for all $x \in \mathcal{D}$, so that

$$((\nabla^{(m+1)}\phi)(x))_{<1>} := (\nabla(\nabla^{(m)}\phi)(x)) \quad \text{for all} \quad x \in \mathcal{D} \tag{22.5}$$

Of course, if ϕ fails to be differentiable, the recursion does not even get started. In general the recursion may break of after a certain $m \in \mathbb{N}^{\times}$ because $\nabla^{(m)}\phi$ fails to be differentiable. We say that ϕ is **of class** C^m if it is m times differentiable and $\nabla^{(m)}\phi$ is continuous. We say that ϕ is **of class** C^{∞} if it is m times differentiable for all $m \in \mathbb{N}^{\times}$.

We often use the notation $\nabla_q^{(m)}\phi := (\nabla^{(m)}\phi)(q)$ to avoid clutter.

Theorem of Symmetry of Higher Gradients. *If the mapping $\phi : \mathcal{D} \longrightarrow \mathcal{D}'$ is of class C^m for a given $m \in \mathbb{N}^{\times}$, then its m^{th} gradient, as given by (22.5), has symmetric values, i.e.*

$$\text{Rng}\,\nabla^{(m)}\phi \;\subset\; \text{Sym}_m(\mathcal{V}^m, \mathcal{V}') \,. \tag{22.6}$$

Proof: We proceed by induction. For $m = 1$, the assertion is trivially valid because of Prop.4 of Sect.02. Assume, then, that the assertion is valid for a given m. Let $x \in \mathcal{D}$ be given. Then, by (22.5) and (22.4), $(\nabla_x^{(m+1)}\phi)_{<1>}$ has symmetric values. Using Def.1 and identifications of the form (02.8), it is easily seen that

$$\nabla_x^{(m+1)}\phi_{<2>} = \nabla_x^{(2)}(\nabla^{(m-1)}\phi) \;\in\; \text{Lin}_{\,2}(\mathcal{V}^2, \text{Lin}_{\,m-1}(\mathcal{V}^{m-1}, \mathcal{V}'))\,,$$

which, by the Theorem on Symmetry of Second Gradients in Sect.611 of Vol.I, is symmetric. Therefore, by Prop.11 of Sect.02, $(\nabla^{(m+1)}\phi)(x)$ is symmetric. ∎

In view of this theorem, we will adjust the codomain of $\nabla^{(m)}\phi$ to $\text{Sym}_{\,m}(\mathcal{V}^m, \mathcal{V}')$ without change of notation, so that (22.3) will be replaced by

$$\nabla^{(m)}\phi : \mathcal{D} \longrightarrow \text{Sym}_{\,m}(\mathcal{V}^m, \mathcal{V}') \tag{22.7}$$

Proposition 1. *Let a family $(\mathcal{V}_i \mid i \in I)$ of linear spaces and a linear space $\mathcal{V}'$ be given and let the given mapping $\mathbf{M} : \times(\mathcal{V}_i \mid i \in I) \longrightarrow \mathcal{V}'$ be multilinear. Then $\mathbf{M}$ is of class C^1 and its gradient is given by*

$$(\nabla_{\mathbf{x}}\mathbf{M})\mathbf{y} = \sum_{j\in I}(\mathbf{M}(\mathbf{x}.j))\mathbf{y}_j \qquad \text{for all} \;\; \mathbf{x}, \mathbf{y} \in \underset{i\in I}{\times} \mathcal{V}_i \,. \tag{22.8}$$

Proof: Recall that the gradient of a given linear mapping is simply a constant whose value is the given linear mapping itself. (See Sect 63 of Vol.I.) Hence the partial gradients $\nabla_{(i)}\mathbf{M}$ are simply the constant mappings whose values are in $\mathrm{Lin}\,(\mathcal{V}_i.\mathcal{V}')$. Therefore, by the Partial Gradient Theorem in Sect.65 of Vol.I, $\mathbf{M}$ is of class C^1, and, by (65.9) of Vol.I, its gradient is given by (22.8). ∎

Now we assume that linear spaces $\mathcal{V}$ and $\mathcal{V}'$ are given

Definition 2. *Given* $m \in \mathbb{N}^\times$ *and* $\mathbf{M} \in \mathrm{Lin}_m(\mathcal{V}^m, \mathcal{V}')$*, we define* $\mathrm{pm}\,\mathbf{M} : \mathcal{V} \longrightarrow \mathcal{V}'$ *, using the abbreviation (02.10), by*

$$\mathrm{pm}\,\mathbf{M} := \mathbf{M} \circ (\mathbf{1}_\mathcal{V}|m)\ , \tag{22.9}$$

We call the resulting mapping a **power mapping of degree** m*. The constant mappings from* $\mathcal{V}$ *to* $\mathcal{V}'$ *will also be called* **power mappings of degree** 0*. The set of all power mapping of degree* $m \in \mathbb{N}$ *obtained in this manner will be denoted by* $\mathrm{Pm}_m(\mathcal{V}, \mathcal{V}')$.

Roughly speaking, the value $(\mathrm{pm}\,\mathbf{M})(\mathbf{v})$ at $\mathbf{v} \in \mathcal{V}$ is obtained by making all the arguments of $\mathbf{M}$ equal to $\mathbf{v}$. Of course, we have, $\mathrm{Pm}_1(\mathcal{V}, \mathcal{V}') = \mathrm{Lin}\,(\mathcal{V}, \mathcal{V}')$. Power mappings of degree 2 will be called **quadratic mappings**, power mappings of degree 3 will be called **cubic mappings**, etc. We have

$$\mathrm{Pm}_2(\mathcal{V}, \mathbb{R}) = \mathrm{Qu}(\mathcal{V})\ , \tag{22.10}$$

i.e., the set of all real-valued mappings of degree 2 is the same as the set of all quadratic forms as defined in Sect.27 of Vol.I.

It is easily seen that the power mapping $\mathrm{pm}\,\mathbf{M}$ remains unchanged if $\mathbf{M}$ is symmetrized according to (03.14) and hence that

$$\mathrm{Pm}_m(\mathcal{V}, \mathcal{V}') = \{\mathrm{pm}\,\mathbf{S} \mid \mathbf{S} \in \mathrm{Sym}_{\,m}(\mathcal{V}^m, \mathcal{V}')\}\ . \tag{22.11}$$

Let $\mathbf{S} \in \mathrm{Sym}_m(\mathcal{V}^m, \mathcal{V}')$ and $j \in (m-1)^]$ be given. Using the notation (02.8) and (02.9), it is clear that

$$\mathbf{S}_{<j>} \in \mathrm{Sym}_j(\mathcal{V}^j\ ,\ \mathrm{Sym}_{m-j}(\mathcal{V}^{m-j}, \mathcal{V}'))\ . \tag{22.12}$$

Applying Def.2 to the above, we find that

$$\mathrm{pm}\,(\mathbf{S}_{<j>}) : \mathcal{V} \longrightarrow \mathrm{Sym}_{m-j}(\mathcal{V}^{m-j}, \mathcal{V}'))\ . \tag{22.13}$$

In the case when j is replaced by 0, we stipulate that

$$\mathbf{S}_{<0>} = \mathrm{pm}\,(\mathbf{S}_{<0>}) := \mathbf{S}_{\mathcal{V}\to\mathrm{Sym}_m(\mathcal{V}^m, \mathcal{V}')} \tag{22.14}$$

the constant mapping on $\mathcal{V}$ with value $\mathbf{S}$.

Proposition 2. *Power mappings are of class C^1 and we have, for every given $m \in \mathbb{N}^\times$ and $\mathbf{S} \in \mathrm{Sym}_m(\mathcal{V}^m, \mathcal{V}')$*

$$\nabla \mathrm{pm}\,(\mathbf{S}) = m \ \mathrm{pm}\,(\mathbf{S}_{<m-1>}) \tag{22.15}$$

where $\mathbf{S}_{<m-1>} : \mathcal{V} \longrightarrow \mathrm{Sym}(\mathcal{V}, \mathcal{V}') = \mathrm{Lin}\,(\mathcal{V}, \mathcal{V}')$ is a special case of (22.10) or (22.12). The gradient $\nabla \mathrm{pm}\,(\mathbf{S}) : \mathcal{V} \longrightarrow \mathrm{Lin}\,(\mathcal{V}, \mathcal{V}')$ is a power mapping of degree $m-1$.

Proof: If $m = 1$, then (22.13) is simply the statement that the gradient of a linear mapping is the constant whose value is the mapping itself. (See p.229 of Vol.I.) Assume that $m > 1$ and let $\mathbf{v}, \mathbf{u} \in \mathcal{V}$ be given. Applying Prop.1 and the chain rule to (22.8), we obtain

$$(\nabla_{\mathbf{v}} \mathrm{pm}\,(\mathbf{S}))\mathbf{u} = \sum_{j \in m^]} (\mathbf{S}(\mathbf{v}|m).j))\mathbf{u}\ . \tag{22.16}$$

By the assumed symmetry of $\mathbf{S}$, we have

$$(\mathbf{S}((\mathbf{v} \mid m).j))\mathbf{u} = \mathbf{S}_{<m-1>}(\mathbf{v} \mid (m-1)).\mathbf{u} = \mathrm{pm}\,(S_{<m-1>}(\mathbf{v}))\mathbf{u}$$

Since $\mathbf{v}$ and $\mathbf{u}$ were arbitrary, (22.16) yields the desired result (22.14). ∎

Proposition 3. *Power mappings are all of class C^∞ and, for every $m \in \mathbb{N}^\times$, we have*

$$\nabla^{(k)} \mathrm{pm}\,(\mathbf{S}) = \left\{ \begin{array}{cl} \frac{m\,!}{(m-k)!} \mathrm{pm}\,(\mathbf{S}_{<m-k>}), & \text{if } k \in m^] \\ 0 & \text{if } k > m. \end{array} \right\} \tag{22.17}$$

for all $m \in \mathbb{N}^\times$ and $\mathbf{S} \in \mathrm{Sym}_{\,m}(\mathcal{V}^m, \mathcal{V}')$.

Proof: Using induction, the formula (22.17) for $k \in m^[$ follows immediately from Prop.2 and Def.1. For $k = m$, the right side of (22.15) is a constant and hence all subsequent gradients are 0. ∎

Corollary 1. *For all $\mathbf{S} \in \mathrm{Sym}_{\,m}(\mathcal{V}^m, \mathcal{V}')$, we have*

$$\nabla^{(m)} \mathrm{pm}\,(\mathbf{S}) = m\,!\,\mathbf{S}_{\mathcal{V} \to \mathrm{Sym}_{\,m}(\mathcal{V}^m, \mathcal{V}')} \tag{22.18}$$

the constant mapping on $\mathcal{V}$ with value $m!\,\mathbf{S}$.

Corollary 2. *For very $m \in \mathbb{N}^\times$ the mapping*

$$\mathrm{pm}\ : \mathrm{Sym}_m(\mathcal{V}^m, \mathcal{V}') \longrightarrow \mathrm{Pm}_m(\mathcal{V}, \mathcal{V}') \tag{22.19}$$

is invertible. In fact, the value of the inverse of pm at $\mathbf{G} \in \mathrm{Pm}_m(\mathcal{V}, \mathcal{V}')$ is the only value of the constant $\frac{1}{m\,!}\nabla^{(m)}\mathbf{G}$.

Definition 3. *Let $\phi : \mathcal{D} \longrightarrow \mathcal{D}'$ be a mapping that is is m times differentiable and let $\nabla^{(m)}\phi : \mathcal{D} \longrightarrow \mathrm{Sym}_{\,m}(\mathcal{V}^m, \mathcal{V}')$ be its m^{th} gradient as determined using* *the Theorem of Symmetry of Higher Gradients. Replacing the values of this gradient by the corresponding power power mapping, we obtain the* m^{th} **power gradient**

$$\nabla^m \phi : \mathcal{D} \longrightarrow \mathrm{Pm}_m(\mathcal{V}, \mathcal{V}') \tag{22.20}$$

defined by by

$$\nabla^m \phi(x) := \mathrm{pm}\, \nabla^{(m)}\phi(x) \quad \text{for all} \quad x \in \mathcal{D} \ . \tag{22.21}$$

In view of Corollary 2 above, the m^{th} gradient can be recovered from the m^{th} power gradient. Therefore, it is often better to deal with power gradients rather than ordinary gradients. Of course, for the first gradient there is no difference, i.e., $\nabla^1 \phi = \nabla \phi$.

the following result is an immediate consequence of Prop 3 and Def.3

Proposition 4. *Given* $m \in \mathbb{N}^\times$, $k \in m^]$, *and* $\mathbf{S} \in \mathrm{Sym}_m(\mathcal{V}^m, \mathcal{V}')$ *the power gradient* $\nabla^{(k)}\mathrm{pm}\,(\mathbf{S}) : \mathcal{V} \longrightarrow \mathrm{Pm}_k(\mathcal{V}, \mathcal{V}')$ *is the power mapping of degree* $m - k$ *given by*

$$\nabla^{(k)}\mathrm{pm}\,(\mathbf{S}) = \frac{m\,!}{(m-k)!}\mathrm{pm}\,(\mathrm{pm}\,(\mathbf{S}_{<m-k>}). \tag{22.17}$$

In particular, the power gradient $\nabla^{(m)}\mathrm{pm}\,(\mathbf{S})$ *is the constant with value* $\mathrm{pm}\,(\mathbf{S})$.

23. Gradients of Power mappings

First, we assume that linear spaces spaces $\mathcal{V}$ and $\mathcal{V}'$, $m \in \mathbb{N}^\times$, and a power mapping $\mathbf{P} \in \mathrm{Pm}_m(\mathcal{V}, \mathcal{V}')$, as defined by Def.2 of Sect,22, are given. In view of Cor.2 of Sect.22, there is exactly one $\mathbf{S} \in \mathrm{Sym}_m(\mathcal{V}^m, \mathcal{V}')$ such that $\mathbf{P} = \mathrm{pm}\,\mathbf{S}$. However, it is often not useful to determine this $\mathbf{S}$ and it is much easier to find a non-symmetric $\mathbf{M} \in \mathrm{Lin}_m(\mathcal{V}^m, \mathcal{V}')$ such that, by Def.2 of Sect.22,

$$\mathbf{P} = \mathrm{pm}\,\mathbf{M} := \mathbf{M} \circ (\mathbf{1}_\mathcal{V}|m) \ , \tag{23.1}$$

We now assume that such an $\mathbf{M}$ has been found. The following formula for the Gradient of $\mathbf{P}$ is easier to use than the formula (22.15) of Prop.2 of Sect. 22.

Proposition 1. *The the gradient of the power mapping* $\mathbf{P}$ *defined by (23.1) is given by by*

$$(\nabla_\mathbf{v}\mathbf{P})\mathbf{u} = \sum_{j \in m^]} \mathbf{M}(((\mathbf{v}|m).j)\mathbf{u}) \quad \text{for all} \;\; \mathbf{u}, \mathbf{v} \in \mathcal{V} \ , \tag{23.2}$$

where the notations (02.10) and (02.4) are used. In words, the j^{th} *term of the sum on right side of (23.2) is* $\mathbf{M}$ *evaluated on the list whose* j^{th} *term is* $\mathbf{u}$ *while all the other terms are* $\mathbf{v}$.

Proof: Put $\mathbf{S} := \mathrm{Str}\ \mathbf{M}$ as defined by (03.14) and let $\mathbf{v}, \mathbf{u} \in \mathcal{W}$ be given. By (22.15) we then have

$$(\nabla_\mathbf{v}\mathbf{P})\mathbf{u} = (\nabla_\mathbf{v}\mathrm{pm}\,(\mathbf{S}))\mathbf{u} = m\ (\mathrm{pm}\,(\mathbf{S}_{<m-1>})(\mathbf{v}))\mathbf{u} \ . \tag{23.3}$$

Using Def.2 of Sect.22 , the definition of $\mathbf{S}_{<m-1>}$) in accord with (03.10), and the abbreviation(02.10), we find

$$(\mathrm{pm}\,(\mathbf{S}_{<m-1>})(\mathbf{v}))\mathbf{u} = (\mathbf{S}_{<m-1>})(\mathbf{v}|m-1))\mathbf{u} = \mathbf{S}(\mathbf{v})|m), m)\mathbf{u})$$

Therefore, using (23.3) and (03.14), we obtain

$$(\nabla_{\mathbf{v}}\mathbf{P})\mathbf{u} = \frac{m}{m!}\sum_{\sigma\in\mathrm{Perm}\,m^{]}} \mathbf{M}(\mathrm{T}_\sigma)(\mathbf{v}|m).m)\mathbf{u})\ . \qquad (23.4)$$

Now let $\sigma \in \mathrm{Perm}\,m^{]}$ be given and put $j := \sigma^{\leftarrow}(m)$. By (01.7), we have

$$(\mathrm{T}_\sigma)(\mathbf{v}|m).m)\mathbf{u}))_i = (\mathbf{v}|m).m)\mathbf{u}))_{\sigma(i)} = \left\{\begin{matrix} \mathbf{v} & \text{if} & \sigma(i)\neq m \\ \mathbf{u} & \text{if} & \sigma(i) = m\end{matrix}\right\}$$

$$= \left\{\begin{matrix} \mathbf{v} & \text{if} & i\neq j \\ \mathbf{u} & \text{if} & i = j\end{matrix}\right\} \qquad \text{for all} \quad i \in m^{]}\ ,$$

and hence that

$$\mathrm{T}_\sigma(\mathbf{v}|m).m)\mathbf{u})) = ((\mathbf{v}|m).j)\mathbf{u}\ .$$

We conclude that, in the sum on right side of (23.4), all the terms for which $\sigma(m) = j$ are the same, namely $\mathbf{M}(\mathbf{v}|m).j)\mathbf{u})$. It is easily seen that there are $(m-1)!$ such terms. It follow that (23.2) is valid. ▮

The result of Prop.1 can be extended to power gradients of higher order as follows:

Proposition 2. *The the k^th power gradient of the power mapping* $\mathbf{P}$ *given by (23.1) is characterized by*

$$(\nabla^{(k)}_{\mathbf{v}}\mathbf{P})\mathbf{u} = \sum_{J\in\mathrm{Fin}_k(m^{]})} \mathbf{M}(((\mathbf{v}|m).J)(\mathbf{u}|J)) \quad \text{for all} \quad \mathbf{u},\mathbf{v}\in\mathcal{V}\ , \qquad (23.5)$$

where the notations (02.10), (02.8) and (02.9) are used. In words, the J-term of the sum on the right side of (23.2) is $\mathbf{M}$ *evaluated on the list whose terms with index in the subset J of $m^{]}$ are all* $\mathbf{u}$ *while the other terms are all* $\mathbf{v}$.

The proof is similar to the proof of Prop.2, and we leave the details to the reader.

We now assume that a linear space $\mathcal{W}$ is given and we consider a variety of power mapping from

$$\mathrm{Lin}\,\mathcal{W}$$

to $\mathbb{R}$ or to $\mathrm{Lin}\,\mathcal{W}$.

The first is the lineonic power $\mathrm{pow}_m : \mathrm{Lin}\,\mathcal{W} \longrightarrow \mathrm{Lin}\,\mathcal{W}\ , m \in \mathbb{N}^\times$, defined in Sect.66 of Vol.I. This lineonic power is the power mapping defined according to (23.1) when $\mathbf{M}$ there becomes the product formation $(\mathbf{L}_i|i \in m) \mapsto \prod_{i=1}^m \mathbf{L}_i$ defined recursively by $\prod_{i=1}^1 \mathbf{L}_i := \mathbf{L}_1$ and

$$\prod_{i=1}^{k+1}\mathbf{L}_i := (\prod_{i=0}^{k}\mathbf{L}_i)\mathbf{L}_{k+1} \quad \text{for all} \quad k\in\mathbb{N}^\times\ , \qquad (23.6)$$

where the product of lineons on the right is described in Sect.18 of Vol.I. For this case, Prop.1 above gives the same result as Prop.4 of Sect.66 of Vol.I, namely

Proposition 3. *The gradient of* pow_m *is given by*

$$(\nabla_{\mathbf{L}}\mathrm{pow}_m)\mathbf{M} = \sum_{k\in m^]} \mathbf{L}^{k-1}\mathbf{M}\mathbf{L}^{m-k} \quad \text{for all} \quad \mathbf{L},\mathbf{M} \in \mathrm{Lin}\,\mathcal{W}\ . \tag{23.7}$$

Prop.2 can be used to obtain the higher power gradients of lineonic power mappings. For example, we have

Proposition 4. *The second power gradient of* pow_m *is given by*

$$(\nabla^{(2)}_{\mathbf{L}}\mathrm{pow}_m)\mathbf{M} = \sum_{\{p,q\in m^]\,|\,p<q\}} \mathbf{L}^{p-1}\mathbf{M}\mathbf{L}^{q-p-1}\mathbf{M}\mathbf{L}^{m-q} \quad \text{for all} \quad \mathbf{L},\mathbf{M} \in \mathrm{Lin}\,\mathcal{W}\ . \tag{23.8}$$

Now put $n := \dim\mathcal{W}$. It turns out the the determinant, the principal invariants, the adjugate, and the principal covariants defined in Chapter 1 are all power mappings. It is easily seen from (14.1) that the determinant function $\det\colon \mathrm{Lin}\,\mathcal{W} \longrightarrow \mathbb{R}$ is a power mapping of degree n, and from (15.3) that the $\mathrm{k}^t h$ principal invariant function inv_k is power mapping of degree k for all $k \in m^]$. Hence, by Prop.3 of Sect.22, all these are of class C^∞.

Proposition 5. *The gradient of determinant function* $\det : \mathrm{Lin}\,\mathcal{W} \longrightarrow \mathbb{R}$ *is given by*

$$(\nabla_{\mathbf{L}} \det)\mathbf{M} = \mathrm{tr}\,(\mathrm{adj}(\mathbf{L})\mathbf{M}) \quad \text{for all} \quad \mathbf{L},\mathbf{M} \in \mathrm{Lin}\,\mathcal{W}\ . \tag{23.9}$$

In the case when $\mathbf{L} \in \mathrm{Lis}\,\mathcal{W}$, *i.e. when* $\mathbf{L}$ *is invertible, (23.9) reduces to*

$$(\nabla_{\mathbf{L}} \det)\mathbf{M} = (\det \mathbf{L})\mathrm{tr}\,(\mathbf{L}^{-1}\mathbf{M}) \quad \text{for all} \quad \mathbf{M} \in \mathrm{Lin}\,\mathcal{W}\ . \tag{23.10}$$

Proof: Let $\mathbf{L} \in \mathrm{Lin}\,\mathcal{W}$, $\mathbf{v} \in \mathcal{W}$, and $\boldsymbol{\lambda} \in \mathcal{W}^*$ be given. By the Theorem on Characterization of Adjugates in Sect 16, i,e,. formula (16.1), and by (25.9) and (26.3) of Vol.I we have

$$\det(\mathbf{L} + s\,(\mathbf{v}\otimes\boldsymbol{\lambda})) - \det(\mathbf{L}) = \boldsymbol{\lambda}\,\mathrm{adj}(\mathbf{L})\mathbf{v} = s\,\mathrm{tr}\,((\mathrm{adj}(\mathbf{L}))(\mathbf{v}\otimes\boldsymbol{\lambda})) \quad \text{for all} \quad s \in \mathbb{R} \tag{23.11}$$

It follows that the directional $(\mathbf{v}\otimes\boldsymbol{\lambda})$-derivative of det at $\mathbf{L}$, as defined by (65.13) of Vol.I, is given by

$$(\mathrm{dd}_{\mathbf{v}\otimes\boldsymbol{\lambda}} \det)(\mathbf{L}) = \mathrm{tr}\,((\mathrm{adj}(\mathbf{L}))(\mathbf{v}\otimes\boldsymbol{\lambda}))\ .$$

Therefore, by Prop. 5 of Sect.65 of Vol.I, we have

$$(\nabla_{\mathbf{L}} \det)(\mathbf{v}\otimes\boldsymbol{\lambda}) = \mathrm{tr}\,((\mathrm{adj}(\mathbf{L}))(\mathbf{v}\otimes\boldsymbol{\lambda}))\ . \tag{23,12}$$

Since $\mathbf{v} \in \mathcal{W}$, and $\boldsymbol{\lambda} \in \mathcal{W}^*$ were arbitrary and since the set of all tensor products $(\mathbf{v}\otimes\boldsymbol{\lambda})$ spans $\mathrm{Lin}\,\mathcal{W}$, the desired result (23.9) follows. The variation (23.10) is a consequence of (16.7). ∎

Proposition 6. *Let $k \in n^{]}$ be given. The gradient of the k^{th} principal invariant function* $\mathrm{inv}_k : \mathrm{Lin}\,\mathcal{W} \longrightarrow \mathbb{R}$ *is given by*

$$(\nabla_{\mathbf{L}}\mathrm{inv}_k)\mathbf{M} = \mathrm{tr}\,((\mathrm{cov}_k(\mathbf{L}))\mathbf{M}) \quad \text{for all} \quad \mathbf{L}, \mathbf{M} \in \mathrm{Lin}\,\mathcal{W} \tag{23.13}$$

The proof is analogous to the proof of Prop.5 above, based on (16.13) rather than (16.1). In view of (16.15) and (15.7), (23.13) actually reduces to (23.9) when $k = n$.

In view of Prop,2 of Sect.22, it follows from Prop.6 above that for each $k \in n^{]}$, the kth principal covariant $\mathrm{cov}_k : \mathrm{Lin}\,\mathcal{W} \longrightarrow \mathrm{Lin}\,\mathcal{W}$ is a power mapping of degree $k-1$. Using Prop.3 and Prop.6 above and also the product rule General Product Rule (see Sect.66 of Vol.I), one can use the formula (16.16) to obtain explicit formulas for the gradients o the principal covariants. The result is the rather complicated formula:

$$(\nabla_{\mathbf{L}}\mathrm{cov}_k)\mathbf{M} = \sum_{j \in (k-1)^{]}} \Big((-1)^{k-1}\mathbf{L}^{j-1}\mathbf{M}\mathbf{L}^{k-1-j} - (-1)^{j}\mathrm{tr}\,((\mathrm{cov}_{k-j}(\mathbf{L}))\mathbf{M})\mathbf{L}^{j-1}$$

$$+(-1)^{j}\mathrm{inv}_{k-j} \sum_{i \in (j-1)^{]}} \mathbf{L}^{i-1}\mathbf{M}\mathbf{L}^{j-1-i}\Big) \quad \text{for all} \ \ \mathbf{L}, \mathbf{M} \in \mathrm{Lin}\,\mathcal{W}\,. \qquad (23,14)$$

Using (16.7), Prop.5 above , and the product rule, we obtain the following formula for the gradient of the adjugate:

$$(\nabla_{\mathbf{L}}\mathrm{adj})\mathbf{M}(\mathbf{L}{+}(\mathrm{adj})(\mathbf{L}))\mathbf{M} = \mathrm{tr}\,(\mathrm{adj}(\mathbf{L})\mathbf{M})\mathbf{1}_{\mathcal{W}} \quad \text{for all} \ \ \mathbf{L}, \mathbf{M} \in \mathrm{Lin}\,\mathcal{W}\,. \tag{23.15}$$

In the case when $\mathbf{L} \in \mathrm{Lis}\,\mathcal{W}$, i.e. when $\mathbf{L}$ is invertible, it follows from (23.15) and (16.7) that the following formula for the gradient of the adjugate is valid:

$$(\nabla_{\mathbf{L}}\mathrm{adj})\mathbf{M} = \det(\mathbf{L})\big(\mathrm{tr}\,(\mathbf{L}^{-1}\mathbf{M})\mathbf{L}^{-1} - \mathbf{L}^{-1}\mathbf{M}\mathbf{L}^{-1}\big) \quad \text{for all} \ \ \mathbf{M} \in \mathrm{Lin}\,\mathcal{W}\,. \tag{23.16}$$

Combining (23.16) with (23.9) we obtain the following result:

Proposition 7. *For every $\mathbf{L} \in \mathrm{Lis}\,\mathcal{W}$, the second power gradient of the determinant function is given by*

$$(\nabla^{(2)}_{\mathbf{L}}\det)(\mathbf{M}) = \det(\mathbf{L})\big((\mathrm{tr}\,(\mathbf{L}^{-1}\mathbf{M}))^2 - \mathrm{tr}\,((\mathbf{L}^{-1}\mathbf{M})^2)\quad \text{for all} \ \ \mathbf{M} \in \mathrm{Lin}\,\mathcal{W}\,. \tag{23.17}$$